Section

Review of Midterm / Sensitivity Analysis

Kentaro Nakamura

GOV 2002

October 31st, 2025

Logistics

- Midterm grade is released
  - Really hard midterm
  - Great job everyone
- We return bluebook
  - Let us know if you have any regrading request
- Today's agenda
  - Review of Midterm Questions
  - Module 7: Sensitivity Analysis

Two Stage Randomized Experiment: Setup

- **First Stage Randomization**: randomizing treated cluster and control cluster
  - W_j: cluster treatment status for cluster j
- **Second Stage Randomization**: randomize treatment for individual within treated cluster
  - $Z_{ij} = 1$: if individual i is in the treated group
  - $Z_{ij} = 0$: if individual i is in the control group

- **Partial Interference Assumption**: No interference between clusters

- **Stratified Interference Assumption**: Individual outcome is affected by (1) their own treatment status and (2) the proportion of the treated units within the same cluster
  - This implies that $Y_{ij} = Y_{ij}(Z_{ij}, W_j)$

Two Stage Randomized Experiment: Estimand

- **Total Effect**: Effect of Treatment + Spillover

$$\tau = \frac{1}{n} \sum_{j=1}^{J} \sum_{i=1}^{n} [Y_{ij}(1,1) - Y_{ij}(0,0)]$$

- **Direct Effect**: Effect of Treatment

$$\delta = \frac{1}{n} \sum_{j=1}^{J} \sum_{i=1}^{n} [Y_{ij}(1,1) - Y_{ij}(0,1)]$$

- **Indirect Effect**: Spillover Effect

$$\xi = \frac{1}{n} \sum_{j=1}^{J} \sum_{i=1}^{n} [Y_{ij}(0,1) - Y_{ij}(0,0)]$$

Two Stage Randomized Experiment: Question 1(c)

- The within-sample variance is given by

$$\mathbb{V}(\hat{\xi} \mid \mathcal{O}_n) = \frac{\sigma_b^2(0,1)}{J_1} + \frac{\sigma_b^2(0,0)}{J_0} - \frac{\sigma_\xi^2}{J} + \frac{1}{m_{01} J J_1}\left(1 - \frac{m_{01}}{m}\right)\sum_{j=1}^{J} \frac{\sigma_j^2(0,1)}{J_1}$$

where

$$\sigma_j^2(z,w) := \frac{1}{m-1}\sum_{i=1}^{m}(Y_{ij}(z,w) - \overline{Y}_j(z,w))^2$$

$$\sigma_b^2(z,w) := \frac{1}{J-1}\sum_{j=1}^{J}(\overline{Y}_j(z,w) - \overline{Y}(z,w))^2$$

$$\sigma_\xi^2 := \frac{1}{J-1}\sum_{j=1}^{J}\{(\overline{Y}_j(0,1) - \overline{Y}_j(0,0)) - (\overline{Y}(0,1) - \overline{Y}(0,0))\}^2.$$

Two Stage Randomized Experiment: Question 1(c)

- Look at the last term

$$\sigma_{\xi}^2 := \frac{1}{J-1}\sum_{j=1}^{J}\{(\overline{Y}_j(0,1)-\overline{Y}_j(0,0))-(\overline{Y}(0,1)-\overline{Y}(0,0))\}^2$$

- This is the sample variance of spillover effect at the cluster level $\overline{Y}_j(0,1)-\overline{Y}_j(0,0)$

- Now, notice that

$$\begin{aligned}\sigma_{\xi}^2 &= \mathrm{var}(\overline{Y}_j(0,1)-\overline{Y}_j(0,0)) \\ &= \mathrm{var}(\overline{Y}_j(0,1)) + \mathrm{var}(\overline{Y}_j(0,0)) - 2\underbrace{\mathrm{cov}(\overline{Y}_j(0,1),\overline{Y}_j(0,0))}_{\text{Not Identified!}}\end{aligned}$$

Two Stage Randomized Experiment: Question 1(d)

- Recall that we have law of total variance

$$\mathbb{V}[\hat{\xi}] = \mathbb{E}[\mathbb{V}(\hat{\xi} \mid \mathcal{O}_n)] + \mathbb{V}[\mathbb{E}(\hat{\xi} \mid \mathcal{O}_n)]$$

- We proved that $\mathbb{E}(\hat{\xi} \mid \mathcal{O}_n)$ is unbiased; thus

$$\begin{aligned}
&\mathbb{V}[\mathbb{E}(\hat{\xi} \mid \mathcal{O}_n)] \\
&= \mathbb{V}\Big[\frac{1}{J}\sum_{j=1}^{J} \underbrace{\Big(\frac{1}{m}\sum_{i=1}^{m} Y_{ij}(0,1) - \frac{1}{m}\sum_{i=1}^{m} Y_{ij}(0,0)\Big)}_{\text{Indirect Effect in Cluster j}}\Big] \\
&= \frac{\mathbb{V}[\overline{Y_j(0,1)} - \overline{Y_j(0,0)}]}{J}.
\end{aligned}$$

- Thus, this part is the variance of indirect effect

Two Stage Randomized Experiment: Question 1(d)

- On the other hand, we already know the form of $\mathbb{V}(\hat{\xi} \mid \mathcal{O}_n)$

- **NOTE**: each σ is not random in finite-population framework, but it is random in super-population framework!

- Therefore, σ should not remain in the last formula
  - You need to take the expectation
  - Fortunately, each σ is unbiased $\rightarrow$ We can replace it with population variance

Two Stage Randomized Experiment with Encouragement: Assumption

- **First Stage Randomization**: randomizing treated cluster and control cluster
  - W_j: cluster treatment status for cluster j
- **Second Stage Randomization**: randomize *encouragement* for individual within treated cluster
  - $Z_{ij} = 1$: if individual i receives encouragement
  - $Z_{ij} = 0$: if individual i does not receive encouragement

- **Assumptions**
  - **Partial Interference Assumption**: No interference between clusters
  - **Monotonicity**: $T_{ij}(z_{ij} = 1, \boldsymbol{z}_{-i,j}) \geq T_{ij}(z_{ij} = 0, \boldsymbol{z}_{-i,j})$
  - **Exclusion Restriction**: $Y_{ij}(\boldsymbol{z}_j, \boldsymbol{t}_j) = Y_{ij}(\boldsymbol{z}'_j, \boldsymbol{t}_j) = Y_{ij}(\boldsymbol{t}_j)$

- We relaxed the stratified interference assumption
  - Thus, $Y_{ij} = Y_{ij}(Z_{ij}, \boldsymbol{Z}_{-ij})$

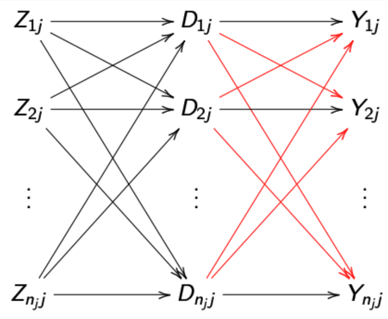
Two Stage Randomized Experiment with Encouragement

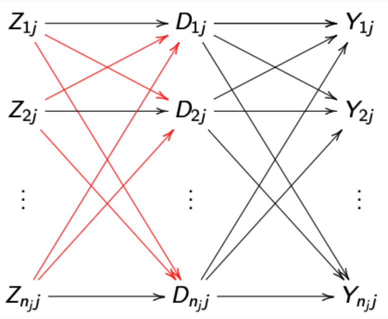
- In question 2(c) and 2(e), we make different assumptions:
  - 2(c): No spillover of treatment receipt on the outcome

$$Y_{ij}(t_{ij}, \boldsymbol{t}_{-i,j}) = Y_{ij}(t_{ij}, \boldsymbol{t}'_{-i,j}) \quad \text{for all } i, j, t_{ij}, \boldsymbol{t}_{-i,j}, \boldsymbol{t}'_{-i,j}.$$

  - 2(d): No spillover effect of encouragement on the treatment receipt

$$T_{ij}(z_{ij}, \boldsymbol{z}_{-i,j}) = T_{ij}(z_{ij}, \boldsymbol{z}'_{-i,j}) \quad \text{for all } i, j, z_{ij}, \boldsymbol{z}_{-i,j}, \boldsymbol{z}'_{-i,j}$$





Two Stage Randomized Experiment with Encouragement

- Numerator is

$$\sum_{j=1}^{J}\sum_{i=1}^{m}\sum_{\boldsymbol{z}_{-i,j}}\{\mathbf{Y}_{ij}(1,\boldsymbol{z}_{-i,j}) - Y_{ij}(0,\boldsymbol{z}_{-i,j})\}$$

$$\times \underbrace{\{T_{ij}(1,\boldsymbol{z}_{-i,j}) - T_{ij}(0,\boldsymbol{z}_{-i,j})\}}_{\text{Only takes 1 for Complier}} \quad \underbrace{\Pr(\boldsymbol{Z}_{-i,j} = \boldsymbol{z}_{-i,j} \mid W_j = 1)}_{\text{Marginalize over Other's Encouragement}}$$

- **Important**: You need to understand when potential outcome is random / not random in finite sample framework
  - Recall that $\boldsymbol{Z}_{ij}$ (encouragement status) is random
  - Treatment is random variable since $T_{ij} = T_{ij}(\boldsymbol{Z}_{ij})$ (similarly $Y_{ij} = Y_{ij}(\boldsymbol{Z}_{ij})$)

- Therefore, even after using consistency

$$Z_{ij}Y_{ij} = Z_{ij}Y_{ij}(Z_{ij} = 1, \boldsymbol{Z}_{-ij})$$

the potential outcome $Y_{ij}(Z_{ij} = 1, \boldsymbol{Z}_{-ij})$ is still random!

Two Stage Randomized Experiment with Encouragement

- Therefore, you cannot do the following!

$$\begin{aligned}
&\frac{1}{J}\sum_{j=1}^{J}\frac{1}{m_1}\sum_{i=1}^{m}\mathbb{E}[Z_{ij}Y_{ij}] \\
&= \frac{1}{J}\sum_{j=1}^{J}\frac{1}{m_1}\sum_{i=1}^{m}\mathbb{E}[Z_{ij}Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-ij})] \\
&\neq \frac{1}{J}\sum_{j=1}^{J}\frac{1}{m_1}\sum_{i=1}^{m}Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-ij})\mathbb{E}[Z_{ij}]
\end{aligned}$$

because $Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-ij})$ is random due to $\boldsymbol{Z}_{-ij}$

Two Stage Randomized Experiment with Encouragement

- Instead, you need to justify as follows!

$$
\begin{aligned}
&\frac{1}{J}\sum_{j=1}^{J}\frac{1}{m_1}\sum_{i=1}^{m}\mathbb{E}[Z_{ij}Y_{ij}] \\
&= \frac{1}{J}\sum_{j=1}^{J}\frac{1}{m_1}\sum_{i=1}^{m}\mathbb{E}[Z_{ij}Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-\boldsymbol{ij}})] \\
&= \frac{1}{J}\sum_{j=1}^{J}\frac{1}{m_1}\sum_{i=1}^{m}\mathbb{E}\left[\mathbb{E}\left(Z_{ij}Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-\boldsymbol{ij}}) \mid \boldsymbol{Z}_{-\boldsymbol{ij}}\right)\right] \\
&= \frac{1}{J}\sum_{j=1}^{J}\frac{1}{m_1}\sum_{i=1}^{m}\mathbb{E}\left[Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-\boldsymbol{ij}})\mathbb{E}\left(Z_{ij} \mid \boldsymbol{Z}_{-\boldsymbol{ij}}\right)\right] \\
&= \frac{1}{J}\frac{1}{m}\sum_{j=1}^{J}\sum_{i=1}^{m}\mathbb{E}\left[Y_{ij}(Z_{ij}=1,\boldsymbol{Z}_{-\boldsymbol{ij}})\right]
\end{aligned}
$$

Sensitivity Analysis

- **Sensitivity Analysis**: Approach to characterize the robustness of your finding

- Two approaches
  - **Approach 1**: Partial R^2 Approach / Omitted Variable Bias Approach
    - Reading: Cinelli and Hazlett (2020, JRSS-B)
  - **Approach 2**: Cornfield Condition (Risk Ratio based approach)
    - Reading: Ding and Vanderwheele (2016, Epidemiology)

- Other approaches: Rosenbaum's Γ
  - Covered in Module 8 (Assuming odds of treatment)

Omitted Variable Bias Formula (1)

- Suppose that true model is

$$Y_i = \alpha + \beta T_i + \boldsymbol{\gamma}^\top \boldsymbol{X}_i + \delta U_i + \epsilon_i$$

but you use the model

$$Y_i = \alpha^* + \beta^* T_i + \boldsymbol{\gamma}^{*\top} \boldsymbol{X}_i + \epsilon_i$$

- Recall that FWL theorem tells us

$$\beta^* = \frac{\mathrm{Cov}(Y_i, \tilde{T}_i^*)}{\mathbb{V}[\tilde{T}_i^*]}$$

where

$$T_i = \phi_0^* + \boldsymbol{\phi}_1^{*\top} \boldsymbol{X}_i + \tilde{T}_i^*$$

Omitted Variable Bias Formula (2)

- Then,

$$\begin{aligned}
\beta^* &= \frac{\mathrm{Cov}(Y_i, \tilde{T}_i^*)}{\mathbb{V}[\tilde{T}_i^*]} \\
&= \frac{\mathrm{Cov}(\alpha + \beta T_i + \boldsymbol{\gamma}^\top \boldsymbol{X}_i + \delta U_i + \epsilon_i, \tilde{T}_i^*)}{\mathbb{V}[\tilde{T}_i^*]} \\
&= \frac{\mathrm{Cov}(\beta T_i + \delta U_i, \tilde{T}_i^*)}{\mathbb{V}[\tilde{T}_i^*]} \\
&= \beta + \delta \times \frac{\mathrm{Cov}(U_i, \tilde{T}_i^*)}{\mathbb{V}[\tilde{T}_i^*]}
\end{aligned}$$

  where the last line is because

$$\mathrm{Cov}(T_i, \tilde{T}_i^*) = \mathrm{Cov}(\phi_0^* + \boldsymbol{\phi}_1^{*\top} \boldsymbol{X}_i + \tilde{T}_i^*, \tilde{T}_i^*) = \mathbb{V}[\tilde{T}_i^*]$$

- Also, consider $U_i = \psi_0^* + \boldsymbol{\psi}_1^\top \boldsymbol{X}_i + \tilde{U}_i$. Then,

$$\mathrm{Cov}(U_i, \tilde{T}_i^*) = \mathrm{Cov}(\psi_0^* + \boldsymbol{\psi}_1^\top \boldsymbol{X}_i + \tilde{U}_i, \tilde{T}_i^*) = \mathrm{Cov}(\tilde{U}_i, \tilde{T}_i^*)$$

Omitted Variable Bias Formula (3)

- Therefore, the bias term is

$$|\beta^* - \beta| = \underbrace{\frac{|\mathrm{Cov}(Y_i^{\perp T, \boldsymbol{X}}, U_i^{\perp T, \boldsymbol{X}})|}{\mathbb{V}[U_i^{\perp T, \boldsymbol{X}}]}}_{=\delta \text{ (By FWL)}} \times \frac{|\mathrm{Cov}(U_i^{\perp \boldsymbol{X}}, T_i^{\perp \boldsymbol{X}})|}{\mathbb{V}[T_i^{\perp \boldsymbol{X}}]}$$

- Now, notice that

$$\frac{|\mathrm{Cov}(Y_i^{\perp T, \boldsymbol{X}}, U_i^{\perp T, \boldsymbol{X}})|}{\mathbb{V}[U_i^{\perp T, \boldsymbol{X}}]} = \underbrace{\frac{|\mathrm{Cov}(Y_i^{\perp T, \boldsymbol{X}}, U_i^{\perp T, \boldsymbol{X}})|}{\sqrt{\mathbb{V}[U_i^{\perp T, \boldsymbol{X}}]}\sqrt{\mathbb{V}[Y_i^{\perp T, \boldsymbol{X}}]}}}_{\text{Partial } R^2 \text{ of } Y \sim U | T, \boldsymbol{X}} \times \frac{\sqrt{\mathbb{V}[Y_i^{\perp T, \boldsymbol{X}}]}}{\sqrt{\mathbb{V}[U_i^{\perp T, \boldsymbol{X}}]}}$$

and

$$\frac{|\mathrm{Cov}(U_i^{\perp \boldsymbol{X}}, T_i^{\perp \boldsymbol{X}})|}{\mathbb{V}[T_i^{\perp \boldsymbol{X}}]} = \underbrace{\frac{|\mathrm{Cov}(U_i^{\perp \boldsymbol{X}}, T_i^{\perp \boldsymbol{X}})|}{\sqrt{\mathbb{V}[T_i^{\perp \boldsymbol{X}}]}\sqrt{\mathbb{V}[U_i^{\perp \boldsymbol{X}}]}}}_{\text{Partial } R^2 \text{ of } T \sim U | \boldsymbol{X}} \times \frac{\sqrt{\mathbb{V}[T_i^{\perp \boldsymbol{X}}]}}{\sqrt{\mathbb{V}[U_i^{\perp \boldsymbol{X}}]}}$$

Partial R-Squared Approach

- As a result,

$$|\beta^* - \beta| = \sqrt{R^2_{Y\sim U|T,X}\frac{\mathbb{V}[Y_i^{\perp T,\boldsymbol{X}}]}{\mathbb{V}[U_i^{\perp T,\boldsymbol{X}}]} \times R^2_{T\sim U|X}\frac{\mathbb{V}[U_i^{\perp \boldsymbol{X}}]}{\mathbb{V}[T_i^{\perp \boldsymbol{X}}]}}$$

- Therefore, with a bit of additional step[1], we get

$$|\beta^* - \beta| = \sqrt{R^2_{Y\sim U|T,X} \times \frac{R^2_{T\sim U|X}}{1 - R^2_{T\sim U|X}} \times \underbrace{\frac{\mathbb{V}[Y_i^{\perp T,\boldsymbol{X}}]}{\mathbb{V}[Y_i^{\perp T}]}}_{\text{Estimatable}}}$$

[1]With FWL theorem, we can indeed derive

$$\frac{\mathbb{V}[U_i^{\perp \boldsymbol{X}}]}{\mathbb{V}[U_i^{\perp T,\boldsymbol{X}}]} = \frac{1}{\frac{\mathbb{V}[U_i^{\perp T,\boldsymbol{X}}]}{\mathbb{V}[U_i^{\perp \boldsymbol{X}}]}} = \frac{1}{1 - R^2_{T\sim U|\boldsymbol{X}}}$$

See Review Question 7 for STAT.

Appendix: Unconditional R^2

- Recall that R^2 for regression $Y \sim \boldsymbol{X}$ is given by

$$R^2_{Y\sim\boldsymbol{X}} = \frac{\mathbb{V}[\hat{Y}_i]}{\mathbb{V}[Y_i]} = 1 - \frac{\mathbb{V}[\hat{\epsilon}]}{\mathbb{V}[Y_i]} = 1 - \frac{\mathbb{V}[Y_i^{\perp\boldsymbol{X}}]}{\mathbb{V}[Y_i]}$$

- Now, notice that since residual $\hat{\epsilon}_i$ is orthogonal to $\hat{Y}_i$, we get

$$\mathrm{Cov}(Y_i, \hat{Y}_i) = \mathrm{Cov}(\hat{Y}_i + \hat{\epsilon}_i, \hat{Y}_i) = \mathrm{Cov}(\hat{Y}_i, \hat{Y}_i) = \mathbb{V}[\hat{Y}_i]$$

- As a result, we can show the connection between unconditional R^2 and correlation coefficient:

$$\mathrm{Cor}(Y_i, \hat{Y}_i) = \frac{\mathrm{Cov}(Y_i, \hat{Y}_i)}{\sqrt{\mathbb{V}[Y_i]}\sqrt{\mathbb{V}[\hat{Y}_i]}} = \sqrt{\frac{\mathbb{V}[\hat{Y}_i]}{\mathbb{V}[Y_i]}} = \sqrt{R^2_{Y\sim\boldsymbol{X}}}$$

Cornfield Condition (Risk Ratio based approach)

- **Setup**: $Y_i(t) \perp\!\!\!\perp T_i \mid U_i$ for $t \in \{0, 1\}$
  - However, U_i is unobserved

- **Estimand**: Now, let's focus on the **causal risk ratio**:

$$RR_{TY}^{\text{true}} = \frac{\mathbb{P}(Y_i(1) = 1)}{\mathbb{P}(Y_i(0) = 1)}$$

  - Risk ratio $= 1$ is equivalent to ATE $= 0$

- We instead observe the observed risk ratio

$$RR_{TY}^{\text{obs}} = \frac{\mathbb{P}(Y_i = 1 \mid T_i = 1)}{\mathbb{P}(Y_i = 1 \mid T_i = 0)}$$

Cornfield Condition (Risk Ratio based approach)

- **Generalized Cornfield Condition**: If $\mathrm{RR}_{TY}^{\mathrm{obs}} > 1$, then

$$RR_{TY}^{\mathrm{true}} \geq \mathrm{RR}_{TY}^{\mathrm{obs}} \times \frac{\mathrm{RR}_{TU} + \mathrm{RR}_{UY} - 1}{\mathrm{RR}_{TU} \times \mathrm{RR}_{UY}}$$

where

$$\mathrm{RR}_{TU} = \frac{\mathbb{P}(U_i = 1 \mid T_i = 1)}{\mathbb{P}(U_i = 1 \mid T_i = 0)}, \quad \mathrm{RR}_{UY} = \frac{\mathbb{P}(Y_i = 1 \mid U_i = 1)}{\mathbb{P}(Y_i = 1 \mid U_i = 0)}$$

Further, in order for $RR_{TY}^{\mathrm{true}} = 1$, we must have

$$\underbrace{\max\{\mathrm{RR}_{UY}, \mathrm{RR}_{TU}\}}_{\text{Unobserved}} \geq \underbrace{\mathrm{RR}_{TY}^{\mathrm{obs}} + \sqrt{\mathrm{RR}_{TY}^{\mathrm{obs}}(\mathrm{RR}_{TY}^{\mathrm{obs}} - 1)}}_{\text{Observed}}$$

Cornfield Condition (Risk Ratio based approach)